European Lighthouse on Secure and safe AI

elsa

DEAR READER,

Mario Fritz © Tobias Ebelshäuser

The *ELSA – European Lighthouse on Secure and Safe AI* is a strong and growing Network of Excellence composed of top experts from research and industry. Our mission is to support the European Union in its efforts to become a beacon of secure and safe artificial intelligence (AI). Key to this mission is connecting emerging and experienced experts in the field of secure AI, addressing the grand challenges in developing secure AI systems, creating common standards for secure AI systems and applications, and promoting young talent.

ELSA began its work in October 2022, building on and expanding the internationally renowned *ELLIS* network (*European Laboratory for Learning and Intelligent Systems*). Together, the 26 founding members of *ELSA* have developed a strategic research agenda, underpinned by targeted research programs focusing on "technical robustness and safety" of AI systems, "privacy-preserving techniques and infrastructures", and "human agency and oversight", which define the broad lines of our research.

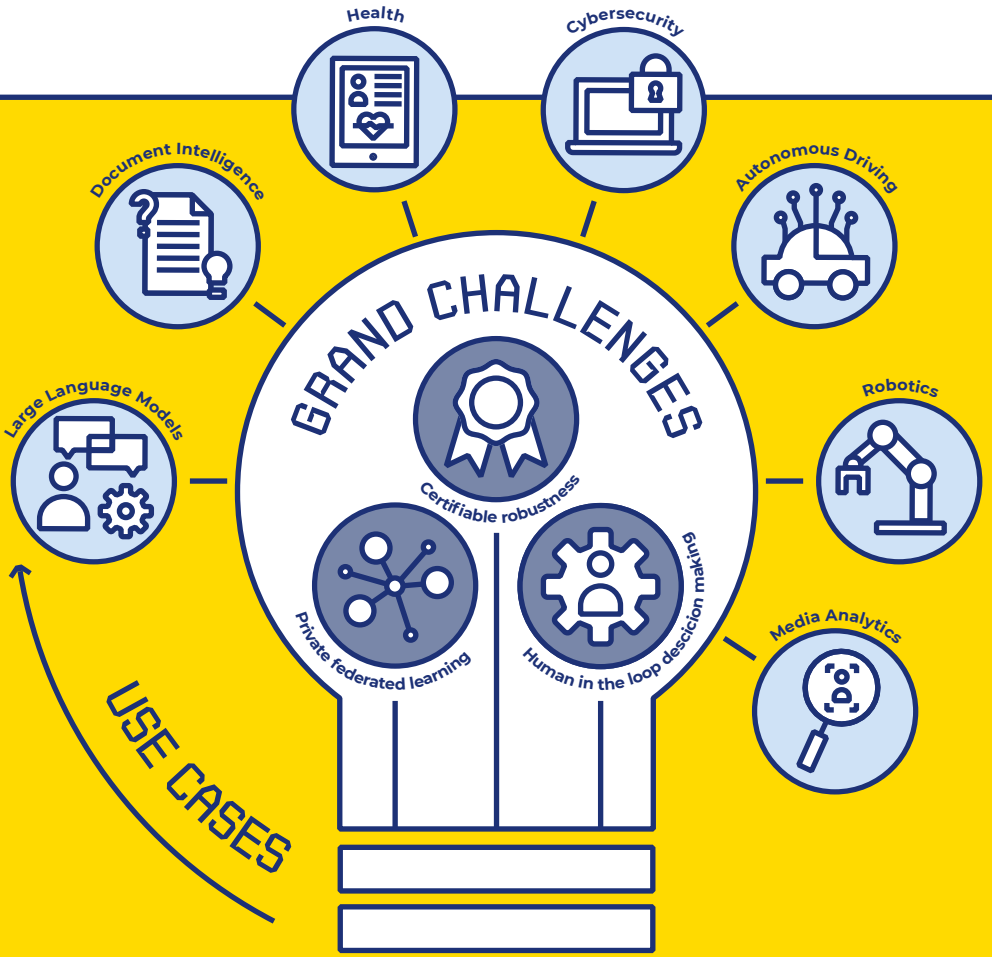
The potentials of AI and machine learning methods are enormous. Our drive is to leverage them for important application fields such as medicine, autonomous driving, robotics, cybersecurity, multimedia, document security, and, last but not least, large language models, while ensuring privacy and security.

This mini-magazine shares how *ELSA* disseminates its knowledge, promotes the exchange between researchers and industry representatives, and ensures the growth of the next generation of excellent researchers in the field of secure AI. Enjoy reading.

Professor Mario Fritz
ELSA-Coordinator and
CISPA-Faculty

FACTS AND FIGURES

ELSA is based on the understanding that safe and secure AI can only be achieved by building on foundational research.



ELSA's progress in terms of methodology and deployment is driven by **grand challenges**. Mastering these challenges is essential in overcoming obstacles to the deployment of AI.

ELSA leverages its broad industrial participation to develop seven ambitious **use cases**. The use cases are central to our methodology and cover a wide spectrum of sectors.

HOW ELSA IS FACING THE GRAND CHALLENGES OF SECURE AND SAFE AI ...

***ELSA's* Strategic Research Agenda**
The increasingly pervasive deployment of AI systems, often based on machine learning, underscores the urgency of enforcing trustworthy AI principles for societal benefit. Our Strategic Research Agenda outlines how *ELSA* will tackle the major challenges of secure and trustworthy AI. *ELSA's* approach includes several cornerstones:

Threat Modelling and Risk Analysis: Our methods are based on rigorous definitions of threats and risks. Once characterized, well-defined statements of properties such as privacy can be given.

Foundational Research and Guarantees: To innovate while adhering to European values, foundational research is key in building trustworthy AI and ML systems. Advances should be grounded in rigorous research to sustain trust in these technologies.

Interdisciplinary Aspect: The success of secure AI technology depends on integrating knowledge far beyond core AI and ML domains.

System View – MLTrustOps: A holistic view of AI/ML systems' design, processing, life-cycle, and impact is essential for security and safety. MLTrustOps includes all relevant aspects in a comprehensive view of AI/ML systems.

Governance and Legal Aspects of Socio-Technical Systems: As AI/ML systems become deeply ingrained in society, governance must ensure compliance and societal wellbeing.

Understanding Limitations and Trade-Offs in Trustworthy AI: While research focuses on addressing pressing challenges, understanding inherent trade-offs and impossibilities is also crucial. This informs technology as well as public discourse, avoiding false promises.

Openness, Transparency, and Accountability: An Open Source approach is vital for transparent and accountable AI development that ensures safety and security.

You can find our SRA here:
https://elsa-ai.eu/sra

... AND HOW YOU CAN GET INVOLVED

The *ELSA* Benchmarks Platform Data sets and benchmarks play a key role in advancing research areas. They unite researchers under a common goal and provide measurable progress. To ensure a targeted approach to innovation, *ELSA* defines benchmarks for its use cases in Autonomous Driving, Document Intelligence, Media Analytics, Health, Robotics, Cybersecurity, and Large Language Models. These benchmarks drive targeted research and allow us to measure progress in addressing major societal and technological challenges.

The *ELSA* Benchmarks platform provides centralized access to all *ELSA* use case-related resources (data, metrics, benchmarks) and allows you to participate in *ELSA* competitions.

Browse through the use cases and find out about current events like competitions, workshops and tutorials. Each use case will organize one or more scientific competitions during the duration of the *ELSA* project. Register on the platform to access the download materials and submit your own methods to the competitions. Moreover, we are always happy to receive feedback.

You can find more Infos here:
https://benchmarks.elsa-ai.eu/

"YOU MEET TOP RESEARCHERS AND BUILD UP A NETWORK."

For early-career researchers, exchanging ideas with top scientists in their research field is crucial. The ELLIS PhD & Postdoc Program connects PhD students and Postdocs from all over Europe with top experts in all areas of machine learning – both in research and industry. The program also enables its participants to spend six months at a partner institution abroad. PhD student Tobias Lorenz has enjoyed the benefits of the ELLIS PhD & Postdoc Program in Oxford. He usually works at the CISPA Helmholtz Centre for Information Security, in the research group of Professor Dr. Mario Fritz, who coordinates the ELSA Network of Excellence.

Tobias Lorenz © Tobias Ebelshäuser

Hello Tobias, you are a participant in the *ELLIS* PhD & Postdoc Program and were in Oxford for an exchange a few months ago. Could you tell me more about the program and what exactly you were up to in Oxford?

The basic idea of the *ELLIS* PhD & Postdoc Program is to strengthen the connections between research groups across Europe. As a participant, I have two supervising professors from two different European countries who give me input on my research work. Part ofthe program is also a stay of at least six months at a partner institution – in my case, this is the University of Oxford. There I worked on my doctoral thesis. My supervisor in Oxford is Professor Marta Kwiatkowska. She is an expert in my research area: the certifiable robustness of machine-learning systems.

Three different tracks are on offer in the *ELLIS* PhD & Postdoc Program: the Academic Track, the Industry Track and the Interdisciplinary Track. You have chosen the Academic Track. Are you planning to stay in the academic world?

At the moment at least, I feel very comfortable in academia because there is so much freedom in science. Also, I find foundational research very exciting. But I could also picture myself working in a research department in industry later on. It remains to be seen what exactly will happen after my PhD.

How does the application process for the *ELLIS* PhD & Postdoc Program work?

There is a central application process via *ELLIS*, in the course of which applicants are matched with professors throughout Europe who are suitable for collaboration on the respective PhD topic. Both sides can then choose who they want to work with. That's how I came to join Mario's group at *CISPA*. I work together with Mario in the *ELSA* Network of Excellence, and I found my second supervisor, Marta, through the *ELLIS* program.

Was working in the *ELSA*-Network a requirement for participation in the *ELLIS* PhD & Postdoc Program?

No, *ELSA* did not even exist back then. Working in one of the networks of excellence might help open doors to the *ELLIS* program, but it is not a requirement. For me working in *ELSA* is very interesting because it involves a lot of research on topics that are very close to my subject. *ELLIS's* scope is much broader and deals with machine learning topics across the board. *ELSA*, on the other hand, focuses on the security and robustness of systems. So, I benefit from both networks.

How exactly did *ELSA* help you to realize your stay in Oxford?

ELSA has a mobility program that is designed precisely for this purpose. I can get a part of my travel expenses reimbursed through this program. My PhD position is also funded by *ELSA*. *ELLIS* itself does not have any positions, but it cooperates universities and research institutions and builds on their programs.

Have you already planned specific collaborations with researchers you met through *ELSA* or *ELLIS*?

There are many ideas already. But I'll have to see how they can be put into practice. Tobias, thank you very much for the interview.

The interview was conducted by Annabelle Theobald.

ELSA supports and expands the pan-European *ELLIS* PhD & Postdoc Program. You can find the full interview and more Information about the Program on our Website:

MORE GOOD NEWS

***ELSA* Mobility Program**
Join the *ELSA* mobility program and explore Europe's vibrant AI research sites. Make contacts, collaborate and gain new insights. Immerse yourself in pioneering projects that make AI safe and secure. The *ELSA* mobility grant offers financial support to PhDs, postdocs as well as experienced researchers if they participate in a research exchange within the *ELLIS* PHD program or the *ELSA* network. Funding is also available for travels to *ELSA*-supported events and workshops.
https://elsa-ai.eu/overview/

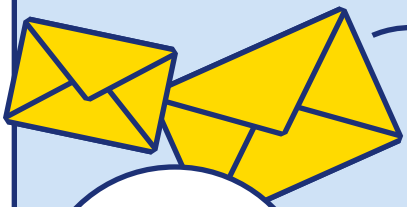
Tejúmádé Àfònjá © Laura Jahke

Oasys Now © Privat

***ELSA* Industry Call**
Issuing our first Industry Call in 2023, *ELSA* called on startups as well as small and medium-sized enterprises to collaborate with us on the development of secure and trustworthy AI applications. After reviewing more than 200 ambitious applications, a jury composed of internal and external experts selected five companies: *Algomo*, *Oasys Now* (shown in the picture), *OttCT*, *QuantPi and Sarus.* Over a period of six months, they will be working together with *ELSA* researchers from across Europe and receive EU funding for their projects. We are thrilled to announce: A second *ELSA* Industry Call has already been issued. Find out more about the winners of the first Call: **https://elsa-ai.eu/winners-call/**

ML Sec Seminar Series
Join us in exploring exciting questions around safe and secure machine learning. The *MLSec Lab*, supported by *ELSA*, is pleased to present a series of seminars with presentations from some of the world's leading researchers in the field. The course program is aimed at machine learning and AI professionals, cybersecurity enthusiasts, researchers and students. Don't miss this unique opportunity to expand your knowledge and skills:
https://elsa-ai.eu/mlsec-seminar-series/

© Laura Jahke



ELSA Newsletter
Do you want to stay up-to-date on our latest research results, next project steps and opportunities to get involved with ELSA? Then sign up for our newsletter now.

IMPRINT

Publisher:
CISPA – Helmholtz-Zentrum
für Informationssicherheit gGmbH
Stuhlsatzenhaus 5
66123 Saarbrücken, Deutschland

Editor-in-Chief:
Annabelle Theobald

Editor:
Annabelle Theobald

Design:
Janine Wichmann-Paulus

Information as of:
August 2024

Photography:
Tobias Ebelshäuser,
Laura Jahke

Contact
ELSA Coordination:
M: elsa-coordination@cispa.de
W: https://elsa-ai.eu

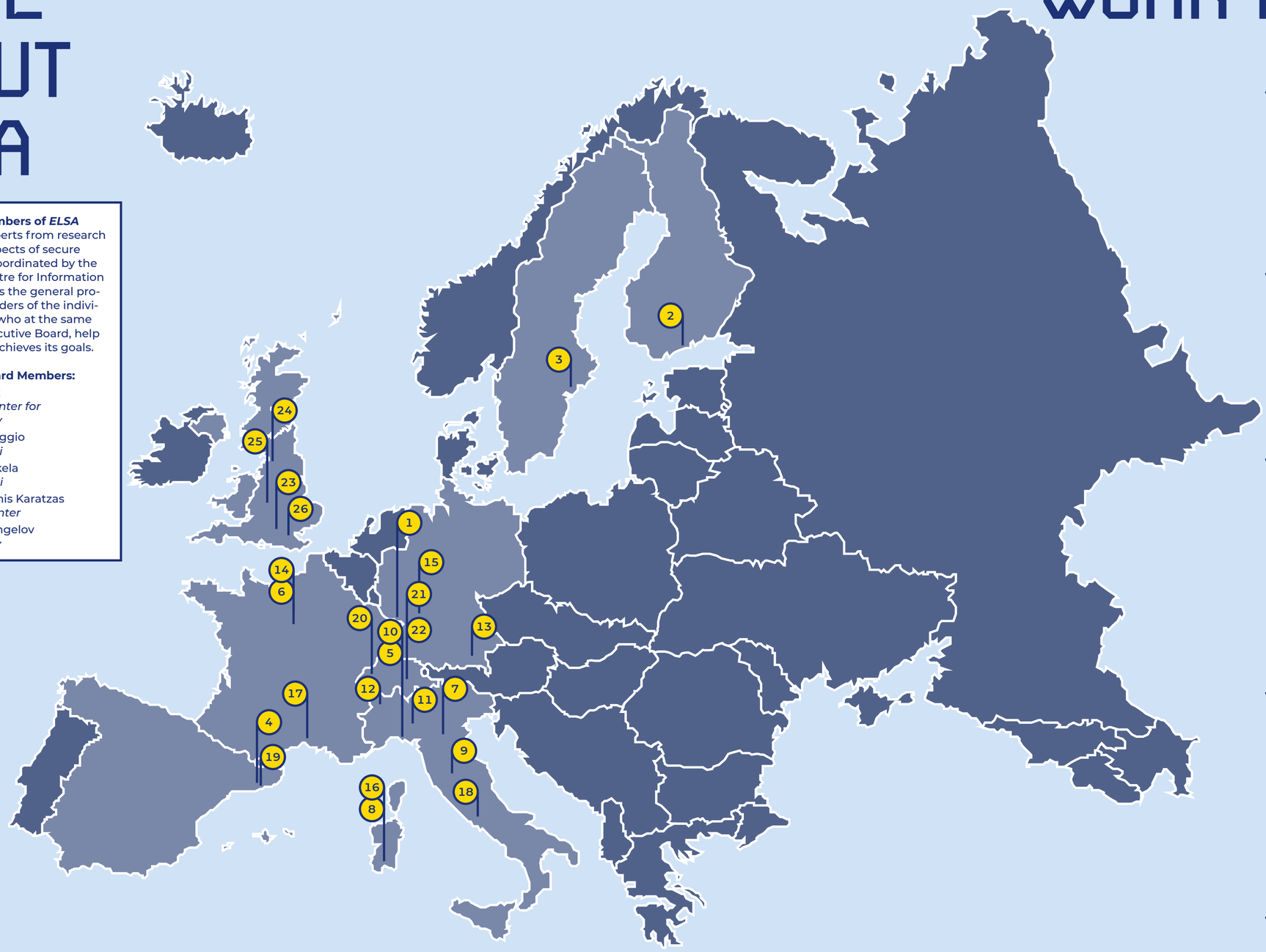
Funded by
the European Union

MORE ABOUT ELSA

The **26 founding members of *ELSA*** include European experts from research and industry in all aspects of secure and safe AI. *ELSA* is coordinated by the *CISPA* Helmholtz Centre for Information Security, which acts as the general project manager. The leaders of the individual work packages, who at the same time form *ELSA's* Executive Board, help to ensure that *ELSA* achieves its goals.

***ELSA's* Executive Board Members:**

- Professor Mario Fritz
  CISPA Helmholtz Center for Information Security
- Professor Battista Biggio
  University of Cagliari
- Professor Antti Honkela
  University of Helsinki
- Professor Dimosthenis Karatzas
  Computer Vision Center
- Professor Plamen Angelov
  Lancaster University



1 *CISPA Helmholtz Center for Information Security*
2 *University of Helsinki*
3 *KTH Royal Institute of Technology*
4 *Computer Vision Center*
5 *Italian Institute of Technology*
6 *Valeo.ai*
7 *University of Modena and Reggio Emilia*
8 *University of Cagliari*
9 *CINI – National Interuniversity Consortium for Informatics*
10 *University of Genoa*
11 *University of Milano*
12 *Polytechnic University of Turin*
13 *Max Planck Society*
14 *National Institute for Research in Computer Science*
15 *European Molecular Biology Laboratory*
16 *Pluribus One*
17 *Yooz*
18 *Leonardo S.p.A.*
19 *PAL Robotics*
20 *EPFL – École polytechnique fédérale de Lausanne*
21 *ETH Zurich*
22 *NVIDIA*
23 *University of Oxford*
24 *Lancaster University*
25 *University of Birmingham*
26 *The Alan Turing Institute*

WORK PACKAGES

Technical Robustness and Safety
Lead: CINI
This work package develops new technologies that enhance the secure use of AI systems. It is led by the *National Interuniversity Consortium for Informatics* (*CINI*). Participating universities include the *University of Cagliari*, the *University of Genoa*, the *University of Milan*, and the *Polytechnic University of Turin*. This team focuses on the concepts of certifiable robustness, resilience, and safety in decision-making processes.

Privacy and Infrastructures
Lead: University of Helsinki
Most modern machine-learning-based AI systems are trained with a huge data set stored in a single centralized database, which limits their applicability. This work package fosters the creation of a new generation of collaborative learning technologies where a number of parties can securely and privately train models for themselves by making use of the combination of data held by all, while being certain that their own data remains secure.

Human Agency and Oversight
Lead: Lancaster University
This work package identifies architectures, mechanisms and methods capable of generating the meaningful and evidence-based assurance that is necessary to secure and maintain the safety and security dimensions of AI systems. Through interdisciplinary investigation between technical experts working in close collaboration with legal, ethics and governance experts, it evaluates existing and emerging methods and mechanisms, and develops new approaches for establishing and certifying secure AI to deliver meaningful and effective AI assurance including human oversight.

ELSA Innovation Lab
Lead: Computer Vision Center – CERCA
This work package tests the technologies and methodologies developed by the other work packages in real-world conditions and prepare assessments of their maturity and applicability. It has established the *ELSA* Innovation Lab: an open, virtual space dedicated to advancing safe and secure AI in a number of application areas defined by the use cases of *ELSA*. The *ELSA* Innovation Lab has established a fluid two-way communication between research activities and real-life needs and applications, both within the network itself and with the wider community.

Communication, Dissemination, Exploitation, Networking
Lead: CISPA
This work package facilitates the development of **ELSA**, which unites the European AI community in the field of safe and secure AI. This team ensures that synergies are utilized and duplication of work is avoided. It also plans how *ELSA's* achievements can be sustainably used and employs targeted communication to ensure the visibility of *ELSA*.

Project Management
Lead: CISPA
This work package covers all aspects of professional project management. The coordinator ensures the proper administrative and financial conduct of the project, supports the coordination of the consortium, and monitors the project progress.